ON EQUIVALENCE OF METHODS OF STATISTICAL MECHANICS

M. DUTTA

A/31, C. I. T. Building, Calcutta-7. India.

(*Received October* 30, 1967 ; *Resubmitted, July* 7, 1968)

ABSTRACT. In statistical mechanics identical results are obtained by methods formulated from widely different considerations. To investigate why and how it is possible, is important. The equivalence of the methods of most probable values, of Gibbs, and of mean values is established here, utilising some recent work in the theory of statistical estimation.

INTRODUCTION

In statistical mechanics, problems are solved by methods, formulated from widely different considerations (Born 1964, Schrödinger 1947, Dutta 1968). All the methods are closely related to the following general methods*:

(i) the method of most probable values, (ii) Gibbs' method with canonical distribution, (iii) the method of mean values.

They yield identical results. A physicist uses one or the other according to his convenience. There are controversies over superiority of one on the others. Born (1964) and Schrödinger (1947) considered them equally useful and reliable but opined slightly in favour of the method (iii). In its favour, Khinchin (1949) criticised the method (i) and (ii) in a way not properly justified. Schrödinger (1947) considered the question of equivalence of these methods as attractive and illuminating, *c.f.*, 'it is a question of a very general theorem of fundamental importance'. Discussions about this question are scattered in the work of physicists but they involve many unnecessary considerations and so are not satisfactory and conclusive from general statistical and mathematical standpoint.

Due to recent work of Dutta (1953, 1955, 1959, 1960, 1965) and Jaynes (1957, 1963) on statistical mechanics from essential statistical (probabilistic) stand-point, it is now possible to discuss the question purely from statistics. In this paper, after a brief report of relevant points from physics and statistics, some salient and significant points about the methods will be just pointed out and this question of equivalence is discussed from the perspective of some recent work in statistics.

*These will be referred to in the text as method (*i*), method (*ii*), method (*iii*).

RELEVANT POINTS FROM PHYSICS

Some casual but significant discussions on this question are seen in the works of Boltzmann (1927) and Gibbs (1900) and in a more systematic way in Ehrenfest and Ehrenfest (1959). Khinchin criticised Gibbs for the introduction of the canonical distribution on ignoring Gibbs' arguments about its sufficiency and without noticing its necessity (Dutta 1966). In different memoirs on the subject including Khinchin's own (1959), this distribution is obtained from different considerations. Schrödinger (1947) and Born (1964) clearly stated the fact of having identical results by different methods but the former ascribed its cause to the insensibility of thermodynamic function and the later to the existence of a very sharp maximum of the distribution function. A good discussion is incorporated in a recent book by Dutta (1967).

RELEVANT FACTS FROM STATISTICS

In the theory of statistical inference from a set of observations on a sample, the form of distribution is guessed in a process of specification and its parameters are determined by methods of the theory of estimation (Fisher, 1938). Of the methods, the following appear to be relevant and important for the present purpose:

(i)′ the method of maximum entropy estimation, (ii)′ the method of maximum likelihood with exponential distribution, (iii)′ the method based on Gauss' principle of arithmetic mean (Kullback 1959, Dutta 1966b).*

In statistics, the method (i)′ is formulated recently (Kullback 1959), following the work of Jaynes (1957). In it, according to Shanon, the entropy of any probability distribution $\{p_j\}$ amongst sample values indexed by j is taken as

$$S = -\Sigma p_j \log p_j \quad \ldots \quad (1)$$

and the distribution is determined by maximising (1) subject to subsidary restrictions

$$\sum_j T_{ij} = \bar{T}_i, \quad i = 1, \ldots, n \quad \ldots \quad (2)$$

when T_{ij}'s are quantities connected with the sample and $\bar{T}$'s, their averages, are supposed to be known.

The basic principle of the method (iii)′ is to determine the parameters by maximising the likelihood function, which Fisher (1938) took as the conditional probability of the sample for a set of values of parameters of distribution of the form, specified suitably. For the present purpose, the law of distribution is taken to be exponential.

The principal of the arithmetic mean was first introduced by Gauss in the theory of observations. Its significance as a method of estimations is due to Keynes (1921) and others (Kendal 1960, Dutta 1966b).

*These methods will be referred to as method (i)′, method (ii)′, method (iii)′ respectively.

Kendal (1960) noticed that the method (iii)′ would yield results obtained by the method (ii)′. Kullback (1959) pointed out that the method (i)′ would yield results obtained by the method (ii)′. The equivalence of the methods (i)′, (ii)′, (iii)′ was established by Dutta (1966b) from general mathematical considerations.

3. ON METHODS (i) AND (i)′

For an assembly of N similar particles corresponding to the energy states $\{\epsilon_j\}$ of a particle, let the occupational numbers be $\{n_j\}$. With the notation, $p_j = n_j/N$, by Boltzmann hypothesis, as usual, one writes

$$S = -k \log w = -k\Sigma p_j \log p_j \tag{3}$$

This is same as the equation (1) except the constant factor k. In the method (i), p_j's are obtained by maximising (3) subject to restrictions of the form (2). Thus, it is the same as the method (i)′.

4. ON THE METHODS (ii) AND (ii)′

Gibbs (1090) introduced two distributions, canonical and microcanonical but he himself considered the latter as a special case of the former. His canonical distribution is

$$P = e^{\frac{\Phi - E}{\Theta}} = \Pi_j e^{\frac{\Phi_j - E_j}{\Theta}} \tag{4}$$

Without following the usual procedure, one may consider (4) as the likelihood function of Fisher for the sample (complexion) and write down the equation for the parameter of distribution, (ψ being the normalising factor), by the principle of maximum likelihood as

$$\frac{\partial \psi}{\partial \Theta} = \frac{\psi - E}{\Theta} \quad \ldots \tag{5}$$

Then, as Born (1964) proposed it, one may identify the right-hand side as the negative entropy. Thus, Gibbs' method with this slight modification is same as the method (ii)′.

5. ON METHODS (iii) AND (iii)′

In the method (iii), the mean values are taken as the normal values for physical quantities. From the probabilistic stand point. the normal values are nothing but most likely (probably) values. Thus, the basis of the method (iii) is a form of the principle of arithmetical mean. Khinchin also accepted this basis and replaced only the mathematical technique of evaluation. Thus, the method (iii) and its modification by Khinchin are basically the method (iii)′. The relations between these different mathematical techniques may be taken as an interesting mathematical investigations.

6. ON EQUIVALENCE

Equivalence of methods (i)′, (ii)′ has been discussed in the section 3. It implies the equivalence of methods (i), (ii) and (iii) as their correspondences are shown in sections 4, 5 and 6. Direct calculations regarding this question of equivalence are not difficult (Dutta, 1968).

CONCLUSION

Identical results, obtained by methods (i), (ii), (iii), in statistical mechanics are not due to something peculiar to thermodynamics or mechanics but are due to equivalence of statistical methods used therein. Even the definition of entropy is intimately related to the likelihood function through the Borel-Canteli central theorem of statistics (Dutta 1966b). Better understanding on these points is expected from deeper studies from probabilistic and statistical considerations.

ACKNOWLEDGEMENT

The author expresses his gratitude to National Professor S. N. Bose F.R.S. for his interest in the problem which was first announced in one of his seminars, and his hearty thanks to Professor Dan E. Christie and Dr. Alan J. Silberger of Bowdoin College, U.S.A. for various help and cooperation.

REFERENCES

Boltzmann, L. 1927, '*Vorlesungen über Gastheorie*', Verlag Johann Ambrosius Berth, 3 Auflage.

Born, Max 1964, '*Natural Philosophy of Cause and Chance*, Dover Publication Inc. New York.

Darwin, C. G. and Fowler, R. H., 1922, *Phil Mag.*, **49**, 450, 823.

Dutta, M. 1953, *Proc. Nat. Inst. Sc. (India)*, **19**, 103, *Bulletine of Int. Stat. Inst.* **32**, 286.

———, 1955, *Proc. Nat. Inst. Sc. (India)*, **21**, 373.

———, 1959, *Proc. Summer Inst. of Theor. Physics of Mussorie (India)*, p. 313.

———, 1960, *Jnan o Bijnan* (in Bengali), **13**, 737.

———, 1965, *Zt. f. Phys. Chem.*, **228**, 380.

———, 1966a, *Indian J. Phys.* **40**, 85.

———, 1966b, *Sankhya*, **28**, 319.

———, 1967, *New Statistical Theory of Gases and Electrolytes in Solution, Scientific Book Agency, Calcutta.*

—, 1968, *Statistical Physics (Foundation)*, World Press, Calcutta (in Press).

Ehrenfest, P. and Ehrenfest T. 1959 : *The Conceptual Foundations of the Statistical Approaches in Mechanics*', Cornell Univ. Press, Y.

Fisher, R. H., 1938, '*The Statistical Theory of Estimation*', Calcutta University Press, Calcutta.

Gibbs, F. W., 1900, '*Elementary Principles of Statistical Mechanics*' Yale Univ. Press, New Haven.

Jaynes, E. T., 1957, *Phys. Rev.*, **106**, 620, **107**, 171.

Kendal, M. G. 1960, *The Advanced Theory of Statistics*, Vol. II, 3rd. Edn., Charles and Griffin & Co., London.

Keynes, J. M., 1921, *A Treatise on Probabilities*', Macmillan and Co. Ltd. London.

Khinchin, A. I., 1949, '*Mathematical Foundations of Statistical Mechanics*'. Dover, Publications Inc., New York.

Koppe, A. H., 1949, '*Statistische Thermodynamik*', S. Hirzel Verlag, Leipzig.

Kullback, S. 1959, '*Statistics and Information*', John Wiley and Sons, New York.

Schrödinger, E., 1947, '*Statistical Thermodynamics*' Cambridge University Press, Cambridge.